Enterprise Integration and Interoperability Improving Business Analytics

Georg Weichhart
PROFACTOR GmbH, Steyr, Austria

Keywords: Enterprise Integration, Enterprise Interoperability, Business Analytics, Smart Grids.

Abstract: In applied research and industrial business analytics (BA) projects data preparation requires around 80% of the total effort. Preparation tasks include establishing technical, semantic interoperability of data and processes to generate value. Enterprise Integration and Interoperability (EI2) approaches address these challenges, but these approaches are hardly taken into account in business analytics. In this position paper, we analyse approaches for their contribution to improving business analytics by supporting the interoperability of data, services, processes and business in general. For more details, we focus on the application domain of smart grids. Existing and missing tool and methodological support as a basis for data-access required for efficient and effective descriptive, predictive and prescriptive business analytics.

1 INTRODUCTION

Before being able to analyse the intersection and contributions of research in Enterprise Integration and Interoperability to Business Analytics, we briefly sketch both fields.

Business Analytics (BA) is a research field where quantitative methods meet decision making and information technology (Hindle et al., 2020). Overall goal is to generate business value by supporting actionable decision making using descriptive, predictive or prescriptive methods. This relies on an appropriate data infrastructure. In addition to this, adaptive organisational capabilities to react to data and results from the methods are required (Dubey et al., 2021; Omar et al., 2019). Business Analytics enables to realize goals like efficiency, effectiveness, flexibility, and resilience. Business and Data Analytics methods are applied to many domains. Electricity Networks is the domain which we will use as application domain below.

Enterprise Integration and Interoperability (EI2) (Weichhart et al., 2021a; Weichhart et al., 2021b) aims to research approaches enabling seamless information exchange between information systems (in the general sense). Multiple systems are enabled to communicate, coordinate and collaborate (Vernadat, 2010) in order to work towards a joint goal. This includes technical systems like software (services), physical systems like networked (manufacturing) machines and social systems like departments. The difference between the two endpoints of a continuum between integration and interoperability, is the degree of coupling. Interoperability aims at loose coupling of systems (e.g. federated interoperability). Interoperability provides approaches for systems-of-systems and Cyber-Physical Systems (CPS), where the systems are developed independently. Integration is focusing on tight coupling, which includes a common information model.

With respect to the underlying assumptions, both, BA and EI2, share some common ground. From a business informatics (in German Wirtschaftsinformatik) point of view, both use IT to support organisational tasks in the context of the enterprise. In this context multiple human actors use different software (and hardware) tools. Efficient and effective support by software for organisational tasks is a key aspect.

In applied research and industrial analytics projects data preparation consumes 80% of the total effort, before any analysis method can be applied. The challenges for data preparation include, identification of data sources, pre-structuring of data sets, cleaning of data, harmonisation and integration of heterogeneous data sets by establishing technical and semantic interoperability. EI2 can fill these gaps. There is a huge potential for EI2 methods and approaches for enabling better analytics projects. In-
teroperability is also needed on business layer with respect to analytics processes to generate value from the data.

In this paper, we discuss, on conceptual level, the suitability of Enterprise Integration, Interoperability approaches for business analytics.

In the following we first discuss and define enterprise integration and interoperability. This is followed by a discussion of data preparation and business analytics challenges. To provide a more concrete picture smart grids (energy networks) have been chosen to provide examples for challenges. In the last but one section, we discuss contributions of EI2 to BA.

2 ENTERPRISE INTEGRATION AND INTEROPERABILITY

We first provide brief definitions of Enterprise Integration and Enterprise Interoperability. We use a model to classify approaches on different levels. These approaches are then discussed.

2.1 The Integration-interoperability Continuum

Enterprise Integration is a research approach, that focuses on the interaction of information systems (in the general sense). Integration of involved systems ensures that an overarching objective is reached by combining the functions of the systems (Morel et al., 2007; Panetto et al., 2012). Systems are analysed on multiple levels. Seamless data and information flows implies that data is exchanged, using a common technical format and semantic information models are defined in a common ontology. All systems and interfaces are aligned to enable this exchange. Services and business processes of organisational systems are aligned to align work-flows to effectively meet business goals.

The above described full and tight integration is one end of the EI2 continuum. In this approach a common model is used by all involved system. This makes it easy to communicate and exchange information, because it works seamlessly and is understood in the same way by all. In short, syntax is standardised, semantics is well defined. Enterprise Integration can be defined as follows: “provide the right information at the right place and at the right time and thereby enable communication between people, machines and computers and their efficient cooperation and coordination” (Kosanke et al., 1999, p.85)

Opposite of tight and full integration, interoperability is found. Enterprise Interoperability is grounded in Enterprise Integration and there is an overlap between approaches. Enterprise Interoperability is extending Enterprise Integration so that it will be possible to meet the requirement of having loose coupled systems (Weichhart et al., 2021a). Interoperability does not impose unified models. Interoperability assumes independent systems that follow their own goals. As such this supports systems-of-systems where systems are heterogeneous and independent. The independence of systems (in a systems-of-systems) supports the evolution of sub-systems. The loose coupling of systems-of-systems (as in contrast to integrated systems), makes them fundamentally different. This is not only true for the structure but also for engineering processes (BKCASE Editorial Board, 2019; Morel et al., 2007). An integrated system may be engineered in a process where parts, which are under full control, are combined. In contrast to that, interoperability is not a static state, but a dynamic (and sometime continuous) process. Individual systems in a system-of-systems maintain their autonomy and evolve. In order to maintain interoperability, the interactions in a system-of-system need to be monitored and interoperability needs to be re-established (by systems-engineers) when it is lost (Ducq et al., 2012; Naudet et al., 2009; Weichhart et al., 2021c).

2.2 Enterprise Integration/Interoperability Frameworks

In the following we are focusing on the problem space described in enterprise integration research (Ducq et al., 2012; Weichhart et al., 2021a). This problem space is very well suited for enterprise integration as well.

The first dimension discussed in EI2 research addresses semiotic levels in the organisation (Stamper, 1993; Stamper et al., 2000). It addresses gaps and barriers between multiple systems. The technological level addresses the encoding of information (syntax). The semantic level addresses the meaning of information. This includes the conceptualisation of information. The organisational level addresses the pragmatics of an organisation. In some frameworks there is also a societal level (legal level). This level includes cultural aspects, encoded as laws, within which the organisation has be behave (Panetto et al., 2019; Weichhart et al., 2021c).

In some approaches a second dimension is used to specify a level of granularity (Chen et al., 2008).
It ranges from data over (technical) services and processes to the business. The service addresses the provision of a single function. The process addresses multiple, coordinated functions. The business the overall enterprise.

The second or third dimension is the solution dimension. It ranges from federated interoperability (loose coupling) to tight integration. In the next section, we discuss approaches that provide interoperability solutions. The approaches are discussed on one level of the problem space, but often address multiple levels.

2.3 Technical Approaches

On technical level, approaches exist that exchange data using messages or service APIs. With service-based approaches both the used API (application programming interface) and the data structures need to be known. Some generic interaction protocols like http(s) provide a more generic standard for coding of web-services building on well known underlying mechanisms like post and get methods to implement a transparent interaction protocol like REST (Representational State Transfer) (Fielding and Taylor, 2002). This particular architectural approach has been designed to support internet scale interoperability.

More recently message-based architectures gained importance. A message broker architecture like MQTT\(^1\) and AMQP\(^2\) provide standardised means to decouple systems. Messages, well known to all participants (on syntax and semantics level), are sent to a broker which provides multiple topics. That component then informs other systems that expressed the interest in the same topic. This decouples processes and supports concurrency.

An older approach to message-based interaction are multi agent systems (MAS) (Wooldridge, 2009). In contrast to broker-based systems there the systems (agents) are communicating directly in a peer-to-peer fashion. The syntax is provided by agent standards like FIPA (Foundation of Intelligent Physical Agents) (FIPA - Foundation for Intelligent Physical Agents, 2005). In that standard (as an example) the semantics of a message is defined through ontologies where in every message the used ontology is specified.

2.4 Semantic Approaches

As mentioned above, ontologies are one way to define the meaning of words and symbols (Guédria and Naudet, 2014; Haller and Polleres, 2020). Ontologies also allow to specify relationships between words and have mechanisms for providing rules. Depending on the goal of an ontology different approaches can be identified. An upper ontology like Cyc (Lenat et al., 2004; Lenat, 2004) defines the meaning of central but often abstract words and symbols to link the semantics of different domains. Middle ontology are more concrete than upper ontology and provide more concrete concepts, but are focused on a domain (of discourse) (Haller and Polleres, 2020). Domain-specific ontologies (e.g. the Ontology of Enterprise Interoperability, OoEI (Naudet et al., 2010)) provide concepts and rules with a particular focus on a domain.

Research in linked data is aiming to publish structured data in a decentralized and bottom-up manner. Linking middle and domain ontologies in a top level ontology an integrated ontology enabling automatic retrieval should be possible and the (Linked Open Data) LOD-Cloud\(^3\) was created (McCrae et al., 2019; Polleres et al., 2020).

This OoEI is one example of a domain specific ontology (defining concepts for enterprise interoperability). Over the years it has evolved, and has been extended using concepts from general systems theory (von Bertalanffy, 1969; von Bertalanffy, 1950). This allows to define the enterprise as a system interacting with other systems in an environment (Guédria and Naudet, 2014).

In another approach, this OoEI has been the basis where the ontology has been replaced by a Domain Specific Language (DSL) (Weichhart et al., 2016a). The goal of that project (OoElf\(^\text{AS}\)) was to re-use the ontological concepts and extend it with an integrated agent model, to support the description of dynamic and complex adaptive systems and their interactions (Holland, 1998). The Domain Specific Language (DSL) in OoElf\(^\text{AS}\) is implemented in the functional programming language SCALA (Wampler and Payne, 2014).

These approaches provide meaning to words and symbols and also provide a conceptualisation of the domain. The used examples above, are from the domain of enterprise interoperability, but several ontologies exist for different domains (Haller and Polleres, 2020).

2.5 Organisational Approaches

Enterprise Interoperability on organisational level is researched by a few approaches. Building on the view that the enterprise is a complex adaptive system (Weichhart et al., 2016b) the S³-Enterprise (Sensing,
Smart and Sustainable) provides different viewpoints on enterprise systems to capture different aspects. The approach follows ISO/IEC 10746 ODP-RM-Open Distributed Processing - Reference Model and defines the following viewpoints (ISO/IEC/JTC1/SC7, 2009):

**enterprise viewpoint:** the enterprise system and its environment

**information viewpoint:** semantics of information and its processing

**computational viewpoint:** functional decomposition and the distribution of (data) objects

**engineering viewpoint:** mechanisms and functions supporting distributed interaction between objects in the system

**technology viewpoint:** choice of technology in that system

Overall goal is an architecture that works like an enterprise operating system (EOS) and services are provided through the EOS abstraction. Data from intelligent sensors is transferred to artificial and human agents for smart decision making (Weichhart et al., 2021c).

The MISA approach is focusing on a methodology for collaborative organisations (Bénaben et al., 2013; Bénaben et al., 2015). It provides a method and supporting tools for interoperability in collaborative, organisational networks. Focus here is also the dynamics in networks and the independence of organisations. This approach also conceptualizes the enterprise as a complex adaptive system.

### 3 DATA PREPARATION IN BUSINESS ANALYTICS

**Business Intelligence** is a term for the processes and tools that allows to discover valuable information and knowledge in the companies’ databases (Zhang et al., 2018). **Big Data** research is driving the ability of systems to handle large and dynamic data (streams). These two research fields provide a technical basis for **Business Analytics (BA)** (Holsapple et al., 2014). However, BA methods include (quantitative) methods rooted in **Operations Research (OR)**, **Machine Learning (ML)**, and **Decision Making** (Hindle et al., 2020).

Overall the rationales for BA is to gain competitive advantage and improved business performance by generating value from data for informed decision making and actionable insight (Holsapple et al., 2014; Hindle et al., 2020).

The current increase in networked information systems and processing power (in the edge and the cloud) reduce the costs and effort for gathering big amounts of data, analysing it with quantitative methods. New insights for business users can be generated in a variety of domains like software, marketing, customer, supply chain, electricity (Holsapple et al., 2014).

Business Analytics methods can be divided in three approaches: (a) descriptive analytics, (b) predictive analytics, (c) prescriptive analytics. The first kind of methods make a situation transparent. The second allow a decision maker (typically using statistical methods) to predict a situation. The last kinds of methods often stem from operations research and quantitative models that allow to analyse multiple courses of action and decision makers are able to make optimized decisions.

“Success in business analytics is a complex matter, depending on a firm’s ability to harness ‘simultaneously’ multiple resources and capabilities (people, process, technology and organization) within a business context, including the data itself (the input and raw material), and deploy these synergistically (key actions and decisions) to deliver a valued output” (Vidgen et al., 2017, p. 634).

According to a study in the smart grid (SG) domain, the most significant barrier in analytics projects is the high data storage and manipulation costs. The second most significant barrier is data complexity and the third data access issues (Bhattarai et al., 2019).

Data management, preparation, manipulation for (business) analytics includes (Vidgen et al., 2017; Zhang et al., 2018):

- integrating heterogeneous data sources
- improving and maintaining data quality
- cleansing and transforming data into common meta-data structures
- governing data and meta-data

These issues are found in many, if not all, analytics project.

### 4 APPLICATION DOMAIN:

**SMART GRIDS**

The Smart Grid (SG) as application domain has many data sources and generates Big Data. Some examples are shown in fig. 1. This figure shows how different data sources in SG are.

For the management of small energy producing units, found typically with sustainable energy generators (wind mills, water turbines, solar energy panels),
many more systems are generating relevant data. In addition to the different types of data sources, these are distributed in a large geographical area.

These sustainable energy systems, dependent on weather conditions are less controllable. Analytics of such energy networks for local micro grids is getting more essential to maintain working energy networks and avoid failures.

Taking the analytics approaches described above we can identify the following services for analytic in the energy domain.

Possible analytics services for SG with special attention to sustainable energy systems includes (Mayilvaganan and Sabitha, 2013; Veerlapati and Thota, 2021; Zhang et al., 2018):

1. Descriptive Analytics
   (a) Asset health monitoring
   (b) Fault detection
   (c) Power quality monitoring
   (d) Detection of energy loss
   (e) Visualization of outage management
   (f) Load disaggregation for reducing energy footprint

2. Predictive Analytics
   (a) Electric device state estimation / health monitoring
   (b) Predictive maintenance
   (c) Condition based maintenance
   (d) Renewable energy forecasting
   (e) Load forecasting and profiling

3. Prescriptive Analytics
   (a) Integrated resource allocation
   (b) Transient stability analysis (resilience analysis)
   (c) Dynamic energy management
   (d) Balance the (predicted) load with energy producers and consumers

The clustering into descriptive, predictive and prescriptive services is not deterministic as most can be placed in multiple categories. For example, load disaggregation is the process of understanding the load different devices generate at customer sites. This supports understanding where energy is consumed the most. However, ultimate goal is to reduce the energy consumption (by prescribing when to turn consuming devices off).

Analysis of stability and the resilience requires to simulate disturbances and analyse a systems possible responses, in order to identify the best response. So it includes a predictive and a descriptive component.

A larger set of standards exists covering many aspects of the smart grid and approaches towards a general architecture have been researched (Uslar et al., 2019), but there are still challenges in the context Business Analytics in the Smart Grid domain. “With the fast deployment of smart meters and advanced sensors, huge amount of data with multiple types and structures from deference sources with a variety of protocols are generated every second. However, the
lack of standard data format for the information software and database structures, as well as the issue of interoperability of different information and communication systems deployed in the smart grids, make it complicated and difficult to obtain data for real application. The traditional way of isolated storage of the data in various systems also increases the barrier for data sharing among applications.” (Zhang et al., 2018, p.19).

Enterprise Integration and Interoperability can provide several approaches that meet the need of BA in general and the SG domain in particular.

5 CONTRIBUTIONS OF EI2 TO BA

Enterprise integration and interoperability (EI2) approaches enable, by definition, a more fluent way of data exchange. Heterogeneous data sources are brought together and the data-structures are homogenized. This capability to make heterogeneous data sources and data models interoperable is a precondition for all analytics approaches. Some EI2 approaches provide direct or indirect support for preparing data for the different types of analytics.

5.1 Descriptive Analytics

EI2 (in general) supports online access to different parts of the system. A traditional data preparation process for business intelligence (BI) is to establish an automated Extraction-Transformation-Load (ETL) process. This process is engineered and leads to a solution where the data source schemata and the targeted business intelligence tool needs to be stable. The fixed schemata allows an automated extraction. To provide an interactive user interface the BI tools need the data in their own - proprietary format. This addresses interoperability on business process, data semantics, and syntax level.

EI2 tools supporting online analytics processes make data sources more transparent and allow ad-hoc queries. This is relevant for distributed systems, where for example supplier specific data is stored locally with the supplier, but access to that data is granted for customers. Also low-code environments for the ETL process support more advanced end-users in getting the right data to the BI and BA tools. An example of such an environment is apache NIFI.

In Smart Grids a number of examples for analytics support that require detailed know-how on the current network state exists. Here EI2 can support BA by providing online access and unification in a common data model for analytics.

5.2 Predictive Analytics

Predictive analytics is one of three general analytics methodological approaches. To cover all three, we mention it here as well, but currently there are no EI2 approaches that specifically support prediction, beyond the basic need for data exchange. However, access to large amount of clean and consistent data is an important precondition for predictive analytics. This includes predictive maintenance for energy assets.

5.3 Prescriptive Analytics

The $S^3$-Enterprise (Sensing, Smart and Sustainable) provides multiple models and points of view to enable an Enterprise Operating System (Weichhart et al., 2016b; Weichhart et al., 2021c). The different model types are used to make processes and data structures in the enterprise transparent. Overall goal is to support smart decision making based on a solid and transparent data basis (Weichhart et al., 2018).

Using process-models for prescription of business behaviour supports interoperability on business level. Modular process approaches like the Subject-Oriented Business Process Management (S-BPM) approach (Fleischmann et al., 2012) supports on the one-hand interoperability in general (Weichhart and Wachholder, 2014) and on the other hand can be used to unify the behaviour of agents so that it becomes interoperable. An example for the later is research in the ROBxTASK project where processes for human and robotic agents have to be aligned (Weichhart et al., 2021d). Such process models can be the result of prescriptive analytics. In particular the ROBxTASK process models will be modular and as such provide task descriptions on a higher level of abstraction that allows decision makers process planning and optimisation without the need for detailed robotic programming know-how (Weichhart et al., 2021d).

To capture dynamic energy processes for management EI2 approaches like the $OoEF^{CAS}$ Domain Specific Language (DSL) can be a basis (Weichhart et al., 2016a). The agent based approach can (by its nature) be easily extended to not only include the systemic aspects of enterprise systems but also of energy systems. This allows to dynamically reason over the current state and possible actions in the network. In addition to this general possibility, does the agent-based approach of $OoEF^{CAS}$ allow to do the analytics tasks in

\footnote{https://nifi.apache.org/}
a distributed manner. That allows immediate reaction to local events. The communication of the agents allows a balancing of energy production and consumption across the network, given that agents have influence over the energy assets they represent.

6 CONCLUSIONS

In this work we have made an initial proposal to align research in Enterprise Integration and Interoperability (EI2) and Business Analytics (BA). The work is motivated by observations that a great share of the effort in analytics projects is for data preparation. EI2 offers existing tools and methods to support this. There is great benefit for BA research to better address the initial steps of the process.

Further work is required to better align tools and methods of both fields. The overall vision is that decision makers, using descriptive, predictive and prescriptive analytics, are enabled to immediately access relevant data also in the context of changing questions. To enable this vision, a specific EI2 approach for meeting business analytics specific needs is to be researched.

Taking a look at all three methodological approaches, initial support for descriptive analytics is already provided by existing approaches. However, end-user tools for immediate change of data sources (e.g. inclusion of CPS and IoT devices) is still a challenge, in particular in the context of distributed systems. We where not able to identify EI2 approaches for prescriptive analytics; more work is needed in this area. While some approaches can be used for predictive analytics, more needs to be done to support end-users in modeling and analysing their current or future system based on interoperable data.

From a more general view, BA currently relies on centralized integrated data sets. This stands in contrast to interoperability and decentralized but interoperable systems. Here, we see an interesting field for further research. How to enable ad-hoc analytics (i.e. reducing setup times and tasks for data preparation) in decentralized systems. For example, smart grids where multiple factories and many households are involved in exchanging energy from sustainable energy sources. Analytics tools are needed for balancing the generation and consumption of power. But all participants have their own point of view and should not have full access to all details in terms of energy needs.

In this initial work we described existing and missing support for business analytics with respect to access to interoperable data sets. More work that supports decision makers in seamless access to data from heterogeneous sources and data processing services is needed.

ACKNOWLEDGEMENTS

The research has been funded by the European Commission and the Government of Upper Austria through the program EFRE / IWB 2020, the priority REACT-EU and the project RESINET (RESIenzsteigerung In energieNETzen – RESilience increase In energy NETworks; Nr.:WI-2020-701900/12).

REFERENCES


